COLLECTING AND ORGANIZING DATA
MEASURES OF CENTRAL TENDENCY
REPRESENTING DATA
Session 2

Topic	Activity Name	Page Number	Related SOL	Activity Sheets	Materials
Collecting and Organizing Data	Collecting Data: Count the Ways	41	K.14, 1.18, 2.23, 3.21, 4.20, 5.18, 6.18, 7.18		Chart paper or transparency, markers
	Random Sampling	43	K.14, 1.18, 2.23, 3.21, 4.20, 5.18, 6.18, 7.18	Data Recording Sheet	Sticky notes, bag or basket
	Household Data	47	K.15, 1.18, 2.23, 3.21, 4.20, 5.18 6.18, 7.18, 8.13	Household Survey Data Sheet	Index cards with data organizer, chart paper, tape, markers
Measures of Central Tendency	Grab A Handful	50	5.19, 6.19, 7.16	Grab a Handful Recording Sheet	Linking cubes, sticky notes, container for cubes
Representing Data	What's Missing?	53	K.15, 1.19, 2.23, 3.22, 4.20, 5.18, 6.18,7.18, 8.12	Candy Graphs A, B, C	
	Object Graphs and Picture Graphs	58	K.15, 1.19, 2.23, 3.22, 4.20, 5.18, 6.18	Ideas for Graphing, T-shirt Pattern	Colored paper squares, linking cubes, scissors, markers, grid paper
	Attributes of Bar Graphs and Attributes of Polygons	62	2.23, 3.21, 3.22, 4.20, 5.18, 6.18, 8.12	Collection of Polygons Sheet, Shading of Polygons Graph	Scissors

Activity: Collecting Data - Count the Ways

Format: Small group

Objectives: Participants will discuss various methods of data collection.

Related SOL: K.14, 1.18, 2.23, 3.21, 4.20, 5.18, 6.18, 7.18

Materials: Chart paper (or transparency) to record ideas, Methods of Collecting Data Activity Sheet

Time Required: 20 minutes

Background: There are a variety of methods for collecting data including counting and tallying, many of which are familiar to children. Children may also have experience with giving or taking surveys. Discussion with students should explore difficulties that can be encountered with surveys. For example, a person may be unwilling to answer a survey/questionnaire/ poll/interview or they may be unwilling to answer the questions accurately. Also, the difficulties that occur in returning and collecting surveys should be discussed.

Children should have an initial understanding of how the above variables can affect the outcome/results. While an interview is typically one-on-one, and the interviewer can somewhat control the conversation, the accuracy of the respondent's answer can still affect the results. Another method of data collection is to examine past records such as polls/surveys conducted, newspaper articles, and public records. However, accuracy of these records cannot be assured, as the method of data collection may be unknown.

Experiments using instruments such as thermometers, yardsticks, calculator-based labs (CBLs), and probes can provide data for use in the classroom.

Simulation can be used to understand natural fluctuations or variation in data. An example is to determine whether a spinner with 6 spaces is "fair". A fair spinner would produce each outcome (1-6) an equal number of times over many, many spins. For instance, if we spun the spinner 60 times, we would expect each outcome to occur approximately 10 times. Notice we say "approximately" because we expect some variation – for instance, maybe eight 1s, thirteen 2s, etc.

Directions:

1. Organize participants into small groups of four (preferably with different grade levels represented) to play a game.
2. Each group will need to select a recorder.
3. Explain to the participants that they will be working in groups to brainstorm as many ways as they can think of to collect data.
4. The catch to this is that they will only get points for their response if no other group in the room has that response. Therefore, they will want to try and think of as many unusual or uncommon methods as possible in addition to methods they believe other groups will think of.
5. Allow groups approximately five minutes to work.
6. When time is up, choose one group to begin the sharing process.
7. As the group leader reads the responses, the instructor will record them on the large chart paper or transparency. After each response, if no other group in the room has that idea written down, the group should record a +1 next to that method. If another group(s) has the method, all groups should cross that idea off of their paper.
8. When the first group is done sharing, the instructor should ask if any other groups have methods that were not read off by the first group. If so, repeat the above process as the next group shares any remaining methods.
9. Sharing continues in this manner until the groups share all possible methods. Some methods that should appear on the list include counting, tallying, measurement, surveys, observations, questionnaires, polls, interviews, examining past records, simulations, and experiments.
10. The group with the most points at the end of this process wins the game.
11. Add to the list any other methods of data collection not thought of by the groups and make the list available during the session to add to by participants as appropriate.
12. Have participants take out their list of grade appropriate stems/questions (from the activities in "Posing Questions", Session I) and identify which of the above methods they would use to gather information to address their stems/questions.
13. Ask them, "When would you use these techniques?" "How would you adapt them to fit the needs of your students?"
14. Have participants copy the group list, or make copies available at end of session.

Activity: Random Sampling

Format: Large Group

Objective: Participants will develop an understanding of appropriate methods of sampling and data collection to ensure that the data provides a representative, unbiased sample of the population.

Related SOL: K.14, 1.18, 2.23, 3.21, 4.20, 5.18, 6.18, 7.18

Materials: Data Recording Sheet, Biased and Unbiased Sampling Methods Sheet, basket or bag from which to draw a sample, sticky notes

Time Required: 45 minutes

Background: As presented in the previous activity, there are several methods for collecting data. Most methods require the researcher to collect data from a population. In most circumstances, collecting information from every member of a population is impossible. Therefore, we collect data from a sample of the population and use the sample to make inferences about the population. Samples can be very accurate in describing the population characteristics. However, for samples to be accurate, they must represent the population. If we wanted to know how much TV middle school students watch, we would not just talk to boys from the eighth grade. This sample of the population only represents one portion of the population—male eighth graders. If the sample is not representative of the entire population, the sample is considered biased because it does not accurately reflect the population being studied.

Two methods that frequently generate biased data are judgment samples and convenience samples. Judgment samples are developed when the researcher uses his/her own judgment to determine what is representative of the population. This method often brings in the researcher's bias about what the population results should be. Convenience samples are developed when researchers select those who are easiest to reach for their sample.

The method used for developing a sample that truly represents the population is random sampling. Randomly selecting a sample does not mean haphazardly selecting members of the population. Rather, it means that each member has an equal

chance of being selected and each sample from the population has an equal chance of being selected. Random methods require that the researcher have an accurate list of the population. The researcher then selects the sample by numbering the list and generating random numbers or putting the names in a hat and drawing out names to be included in the sample. When generating the sample, sample size needs to be considered. Enough data needs to be collected to be sure that conclusions are accurate.

A key point to make through this activity is that all samples vary. Rarely will a researcher take two samples from the population and get the same results. Some students think of sample variation as bias. Bias is difficult to judge from one sample. Sampling methods that are biased will show their bias over several samples. For instance, we cannot prove that a coin is an unfair coin by flipping it once or even ten times. We might get more than five heads if we flip a coin only ten times. However, if we continue to flip the coin ten times and our results consistently show more than five heads, we can say the coin is biased. Over several trials, random, unbiased samples will represent the population.

Directions:

1. In this activity, participants examine the question "What is the average number of years of teaching experience for participants?" The activity requires that the participants develop three different sample representations of the data: 1) judgment sample, 2) convenience sample, and 3) random sample.

2. Create several **judgment samples.** Ask the participants to make a prediction about the average number of years of teaching experience. To create this sample, have them ask five participants who they deem representative of the population in the room. Have the participants put their average, rounded to the whole number, on a sticky note and build a "sticky note line plot" on the board. To build the line plot, draw a horizontal line on the board and mark a scale on it in whole number units. Have participants put the sticky notes on the board above the scale mark that goes along with their average, stacking repeat numbers. Ask the participants to comment on what they see with these sample averages, prompting them with the following questions:
   - Did everyone get the same sample average?
   - How much variation is there in the averages?
   - What would your prediction of the population average be?
   - Do you feel very comfortable with your prediction? Why?

3. Create several **convenience samples.** Ask the participants to take a convenience sample of the population by selecting five participants near them either sitting at their table or close by. Have the participants put their average on a sticky note and build a "sticky note line plot" on the board. Again ask the participants to comment on what they see with these sample averages. Ask the following questions:
   - Did everyone get the same sample average?
   - How much variation is there in the averages?
   - What would your prediction of the population average be?
   - Do you feel very comfortable with your prediction? Why?

4. Compare the two line-plots created in section 2 and 3.

5. Create several **random samples.** Collect the data from the participants in the room. As you do this, have one person at each table collect the data on the data sheet. Have the participants cut out the data into small squares and place them into a bag or basket. Have each participant take a sample of five squares from the data. The squares should be replaced before another participant takes a sample. Have the participants calculate the average for their sample and create a third "sticky note line plot." Examine the line plot, determining its mean, and comparing it to the mean and variation of the other two line plots. Ask participants to comment on the differences in predictions one would make from the three types of samples. Overall, we should see that the distribution of sample averages from the random sample will be less varied than the other samples and should be an unbiased representation of the data.

6. Find the true average for the population of participants in the room. Compare this average to the averages developed from each of the sampling methods. If all works well, the random samples overall should be the most accurate. However, samples will vary and therefore, some random samples will not be as good as some of the judgment samples. In addition, the samples are small; therefore more variation should be expected. It is possible that the random sample does not provide the most accurate estimate of the true average. If there is time to discuss the concept of sample size, it would be helpful to repeat the random sampling process but have the participants take samples of ten rather than five squares and collect the average data once again.

Data Recording Sheet

Activity: Household Data

Format: Pairs/Small group

Objectives: Participants will organize a set of data using one of the data organizers discussed in the session. Pairs will share what their data looks like using the appropriate organizing tool.

Related SOL: K.15, 1.18, 2.23, 3.21, 4.20, 5.18, 6.18, 7.18, 8.13

Materials: Chart paper, markers/pens, tape, Household Survey Data Activity Sheet
Prepare index cards beforehand with the following data organizer directions (one direction per card):

1. Use a list to show the number of adults in each household.
2. Use a chart to show the ratio of cars to adults.
3. Use a chart to show the ratio of TVs to people.
4. Use a frequency table to show the number of people in households.
5. Use a stem-and-leaf plot to show the number of TVs in households.
6. Use a box-and-whisker plot to show the number of children in households.
7. Use a frequency table to show the number of cars in households.
8. Use ordered pairs to show the relationship between the children in a household and the total number of people in a household.
9. Use a matrix to show the number of children versus the number of cars per household.
10. Using a tree diagram, show the total number of combinations of children and cars.

Time Required: 30 minutes

Directions:

1. Organize participants into pairs. Explain to the participants that they will be working together to organize a set of data.
2. Brainstorm together ways that data can be organized. Make a chart or write ideas down on a transparency.
3. When all methods are thought of/discussed, give each pair an index card with a data organizer written on it. Give all pairs the same Household Survey Data Activity Sheet.

4. Pairs need to use the data organizer from their index card to organize the Household Survey data they just received onto chart paper.
5. Have pairs post their displays on the wall for others to see.
6. If time, have pairs share the similarities/differences they notice about the different ways the data is displayed.
7. Discuss why the same set of data looks different. Does the organization of the data change its meaning?

1999 Survey of Household Members

Name	Children in Household	Total Number of People in Household	Number of TVs in Household	Number of Cars in Household
Adams	2	4	4	2
Brown	1	3	2	2
Bury	0	3	3	3
Chambers	4	6	5	3
Cleveland	1	3	2	2
Critzen	2	3	3	1
Cunningham	2	4	2	1
Davis	3	5	4	3
Dumante	1	2	2	2
Elliot	2	3	3	3
Gale	0	1	2	1
Galland	0	2	3	2
Herrig	2	4	4	2
Kerby	1	3	2	2
Kincaid	2	4	2	2
Leigh	2	3	3	1
Lowe	1	3	2	2
Mazick	2	4	3	2
Martinez	3	5	5	3
Moore	2	3	4	2
Nunez	2	5	4	3
Pranter	2	4	4	1
Richards	0	2	1	1
Riley	0	1	3	1
Roberts	1	3	3	2
Shaw	2	3	3	1
Whitten	1	2	2	1

Activity: Grab a Handful

Format: Large Group

Objective: Participants will collect data and use the data to illustrate measures of central tendency and range.

Related SOL: 5.19, 6.19, 7.16

Materials: Container of linking cubes, sticky notes

Time Required: 30 minutes

Directions

1. Participants grab a handful of cubes from a bucket or bag. The handful should be as large as possible. (At this point the leader should take only two cubes to be used in the discussion of *outliers* in #5.)
2. Each participant connects all of his cubes and writes the number of cubes in his train on a sticky note.
3. Participants then take their sticky notes to the chalkboard and construct a bar graph or a line plot.
4. Participants then line up at the front of the room according to the number of cubes each participant has. They should line up in front of each other when there is more than one participant with the same number of cubes. This human representation of the data should be similar to the line plot or bar graph on the board with sticky notes.
5. The instructor leads the group through a discussion of the measures of central tendency using the vocabulary words. While discussing the range, participants learn about outliers. Explain that an outlier is a value widely separated from other values in the data set. The two cubes that the leader picked in #1 above represent an outlier, as this value should be far separated from the value of the smallest handful.
6. Before discussing the words mode, median and mean, participants should be asked what they know. Assessing prior knowledge is key. They also should be encouraged to predict the median and the mean. The mode is easily illustrated on the human representation of the data and the line plot on the board.
7. Before the discussion of the median begins, ask participants to spread out in one line, maintaining the sequential order of the data from least to greatest.
8. Ask participants to move out from the line in pairs – one from each end at the same time. This allows participants to see how, as you approach the

center, one or two participants will remain. They experience the meaning of the center. The movement away from the center makes this very clear to participants. Discussion should follow regarding the median and what it represents.

9. Before illustrating the mean, participants should discuss again what they are really trying to find. Encourage participants to share ways they could find the mean using the cubes.
10. Have participants share linking cubes with each other trying for all participants to have the same number of cubes. They share until sharing anymore isn't helpful to reach the same number of cubes for each participant. More than likely two groups will exist. Some participants will have trains of one number and the other group will have trains of another number. At this time participants discuss the mean and also use the term average.
11. End the session by reviewing the measures of central tendency and stating again the actual answers for this collection of data. Participants should be encouraged to discuss how changes in the data would affect the measures of central tendency. Participants should explore the results of changing the data near the center as well as at the extremes. Have participants compare their results and describe the changes that occur. Ask participants to complete the Recording Sheet.

Grab a Handful

Recording Sheet

Name_____
Date_____

Number of cubes I grabbed in one handful_____

Number of participants participating in this activity_____

Record the number of cubes collected by each participant on the back of this sheet.

Using the data, I discovered the following:

Range of the data _____

Identify any outliers, and tell why they are outliers.

Mode(s) of the data _____

How did you know?

Median of the data _____

What is the meaning of median?

Mean of the data _____

Why is the mean also called the average?

Which measure of central tendency best represents the data and why?

Activity: What's Missing?

Format: Whole Group

Objectives: Participants will learn the different parts of a graph and the importance of each part.

Related SOL: K.15, 1.19, 2.23, 3.22, 4.20, 5.18, 6.18, 7.18, 8.12

Materials: Activity Sheets of graphs of survey results about candy, Candy Graph A, B, and C

Time Required: 10 minutes

Background: As we begin to teach data representation methods to children, it is important to discuss the "big picture" of the data. What is this data representing? How do you know? What information on the graph tells you what you need to know? What is the sample size? What is the scale?

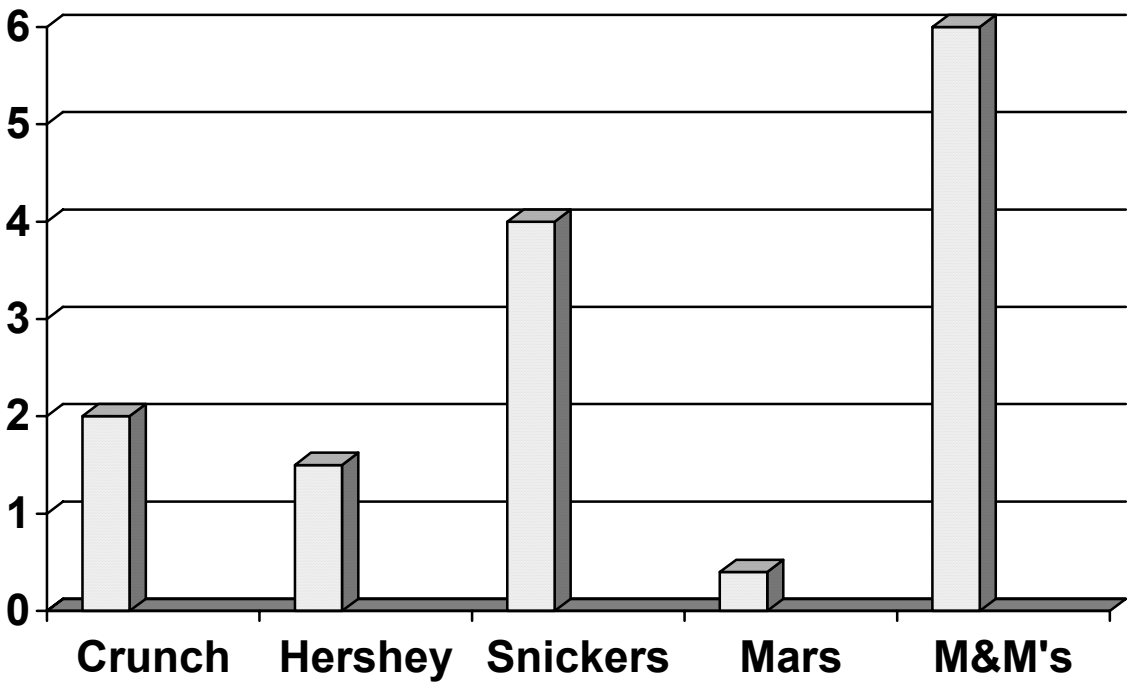
Oftentimes, children don't understand that each unit on the graph is of equal size and, therefore, represents an equal amount. It is the number of units that shows the amount one is talking about. A ***scale*** is a system of marks in a given order and at specific intervals. The interval should be constant. In most graphs, each ***axis***, or reference line, should be labeled to show the categories being represented as well as the scale. If a picture or symbol is being used, a ***key*** should be included which states the unit value of the picture or symbol. The ***title*** should indicate the information the graph represents. Sometimes a ***sample size***, or number of people surveyed, is included, but can also be determined from some graphs by totaling the number of objects represented.

Directions:

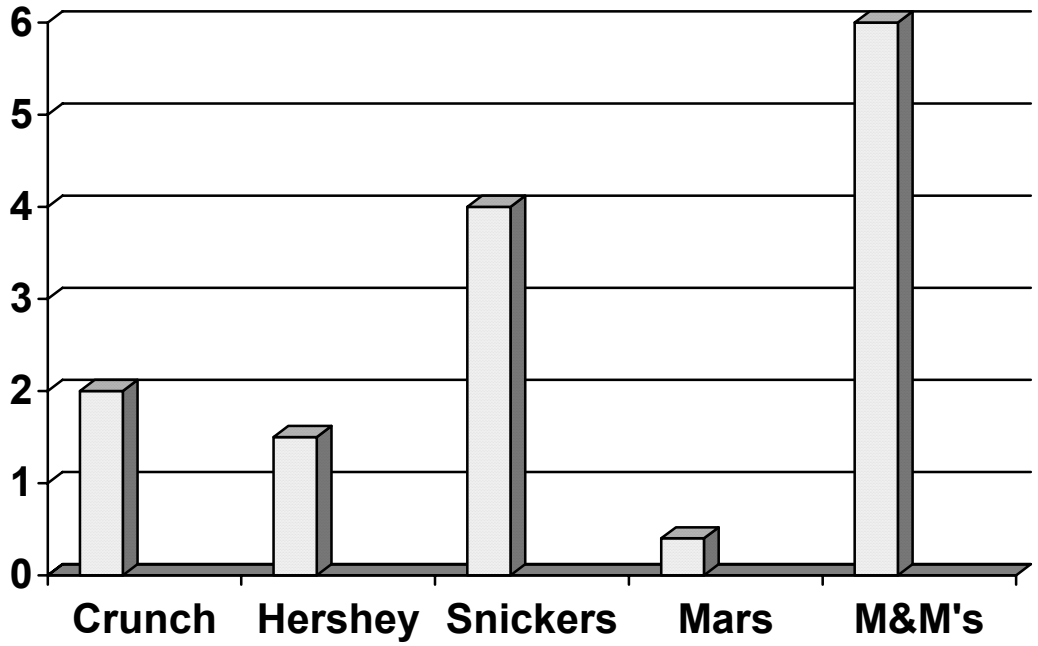
1. Display Candy Graph A.
2. Ask participants what they think this information represents. What would be a good title for this graph? How would you label the axes? Write down suggestions.
3. Display Candy Graph B.
4. Repeat the above questions. What is different about this graph? How does this new piece of information change the way one interprets this graph?

5. Display Candy Graph C and discuss all parts of the graph including labels, axes, title, scale, etc.

Candy Graph A

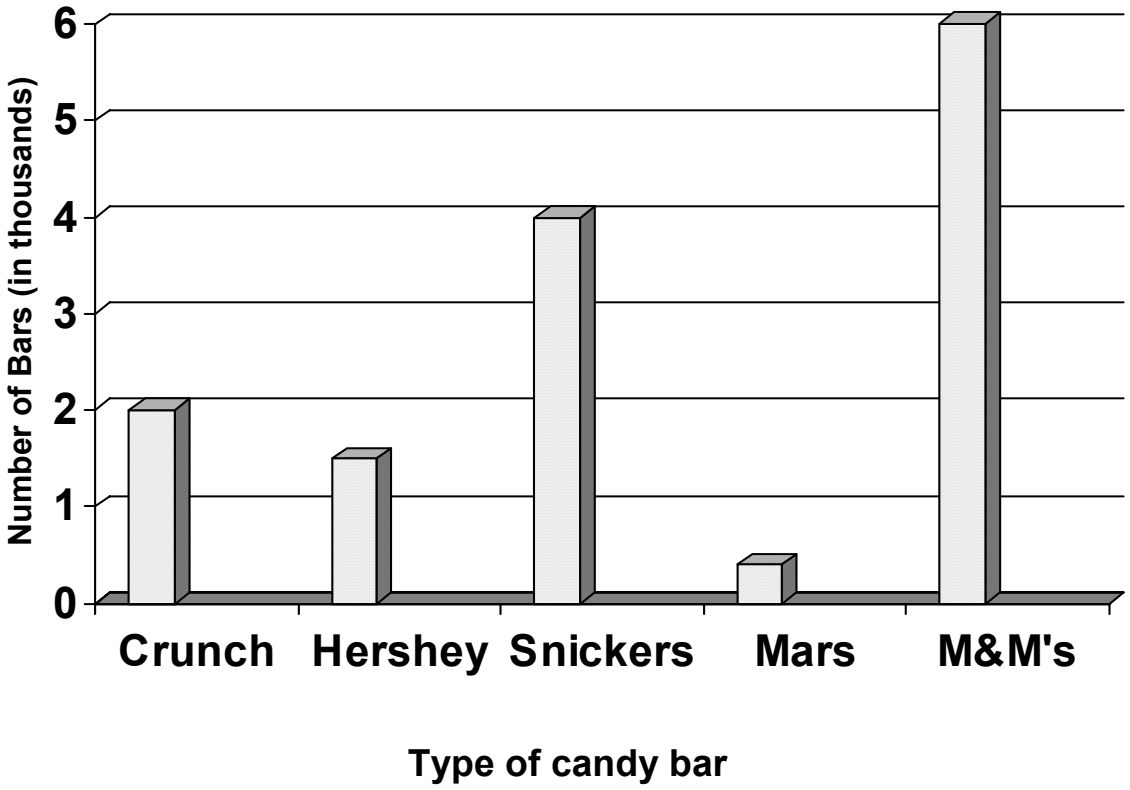


Candy Graph B



Candy Graph C

Number of Candy Bars Sold in the Cafeteria in 1999



Activity: Object Graphs and Picture Graphs

Format: Whole Group

Objectives: Participants will learn how to construct a pictograph and a bar graph.

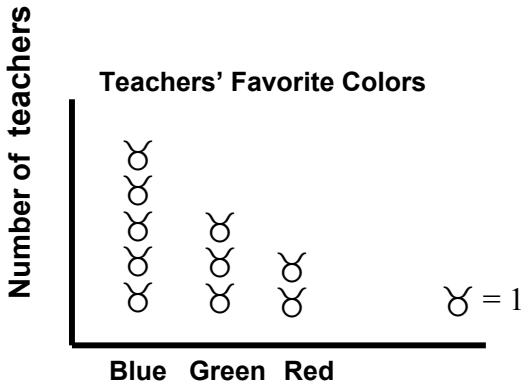
Related SOL: K.15, 1.19, 2.23, 3.22, 4.20, 5.18, 6.18

Materials: Multilink or unifix cubes of various colors, construction paper squares to match the colors of the cubes, t-shirt pattern, scissors, crayons or markers, grid paper

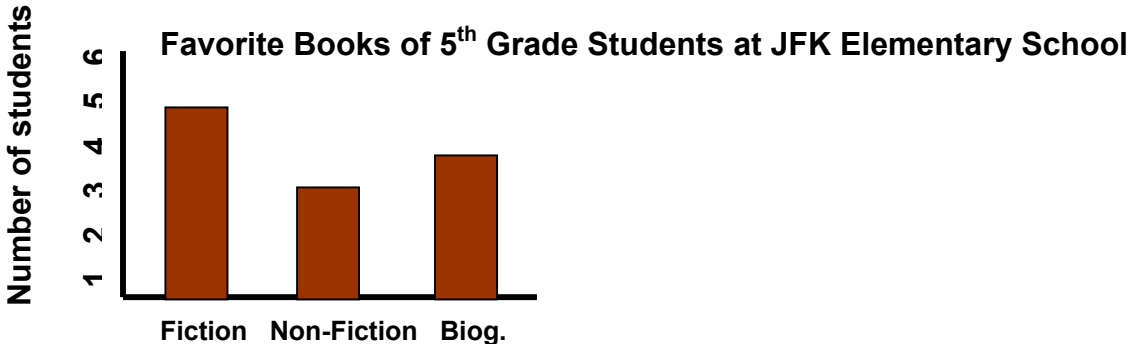
Time Required: 30 minutes

Background: An ***object graph*** uses a three dimensional object to represent a unit.

A ***picture graph*** uses a picture or symbol to represent an object. If there is more than one of the objects, multiple representations of the symbol are used. A key should be included that states the value of the symbol.



A ***bar graph*** differs from a pictograph in that either horizontal or vertical parallel bars are used to represent counts of data in several categories. A bar is used for each category with the length of the bar representing the total count for that category.



Directions:

Object Graph

1. Place a handful or two of colored cubes on each table.
2. Ask participants to take a cube that most closely matches the color of their clothing.
3. Collect the data by having participants place their cubes on the matching colored paper square that the instructor has placed around the room.
4. Connect all of the cubes on the colored squares and make an object graph with the cubes. Place them on the floor or tape them on the wall.
5. Add a title, label the axes, and create a key.

Picture Graph

1. Next, have participants color in the t-shirt pattern to match their cube and cut it out.
2. Ask for suggestions on how to organize the data (sort by colors).
3. Tape each t-shirt up on the wall to construct a pictograph with appropriate labels and key.
4. Ask participants for the differences they notice between the two graphs constructed so far.

Bar Graph

1. Give each participant a piece of grid paper.
2. Have them construct a bar graph from the t-shirt color data.
3. Discuss differences between the three types of graphs.
4. As a large group, discuss other ideas for data collection appropriate in a classroom. Distribute copies of “Ideas for Graphing.”

Ideas for Graphing (adapted from Marcy Cook)

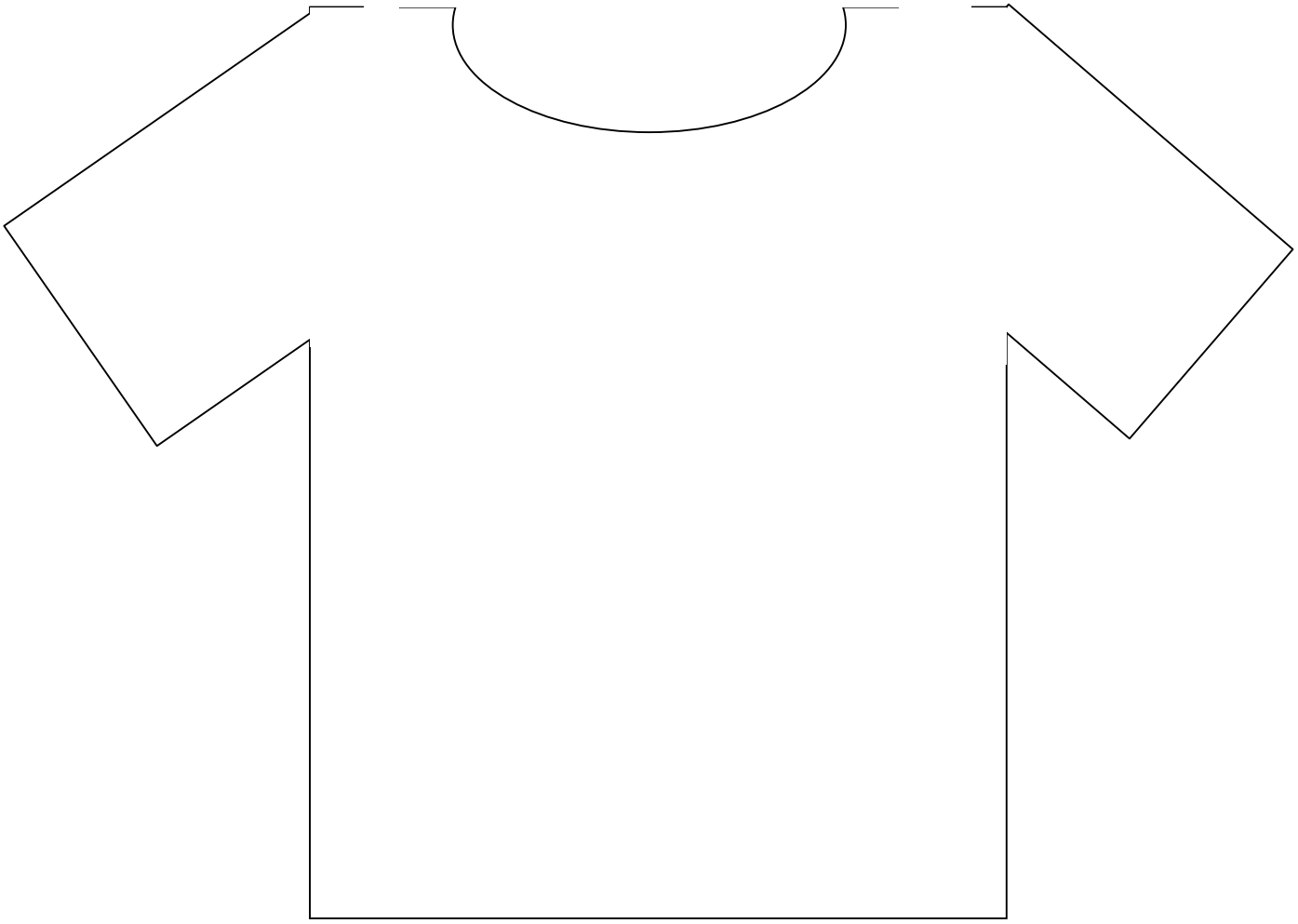
Bar Graphs

- In which quarter of the year is your birthday?
- Is the date of your birthday even or odd?
- Estimate how many times you went swimming last summer.
- What is the sum of the digits in your address?
- What is the difference between the largest digit in your address and the smallest digit in your address?
- How many objects in your desk have a number on them?
- How many doorways can you walk through in your house?
- What is the sum of the digits in your telephone number?
- How many times do your eyes blink in one minute?
- What is your favorite subject in school?
- What subject do you think we should spend the most time on?
- Write a complete sentence. Which letter did you use the most?
- How many times can you write the number 8 in one minute?
- How many different words ending with the letter "E" can you write in one minute?
- Approximately how many times do you get out of your seat on a school day?
- What time do you wake up on a school day (to the nearest quarter hour)?
- How many different vowels are used when you write your full name?
- Determine the fractional part of your name that is consonants.
- Collection of Halloween candy.

Line Graph

- Keep track of your bedtime for one week.
- Keep track of how many glasses of water you drink for one week.
- Collect the daily temperature for one month.
- Attendance (class or school) for a week, month, etc.
- Stock prices for a week, month, etc.
- Money collected each hour during a school fundraiser (car wash, bake sale, etc.).
- Number of visitors to your house every ten minutes on Halloween.
- Amount of money collected by the lunchroom each day for one week.
- Number of math problems assigned for homework each day for one week, month, etc.
- Length of time it takes you to get to school each day for one week.
- Amount of time you spend on homework each night for one week, month, etc.

T-shirt pattern



Activity: Attributes of Bar Graphs and Attributes of Polygons

Format: Large Group/Small Group

Objective: Participants discover the attributes of well-constructed bar graphs.

Related SOLs: 2.23, 3.21, 3.22, 4.20, 5.18, 6.18, 8.12

Materials: Collection of Polygons Activity Sheet, paper, rulers, colored markers

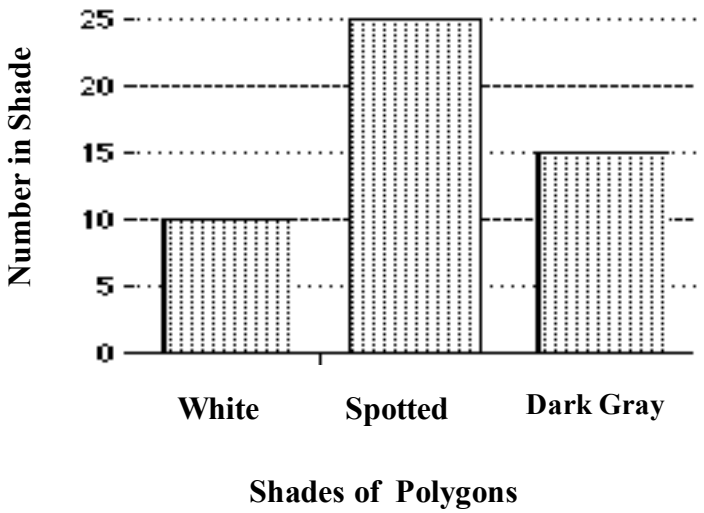
Time Required: 30 minutes

Directions:

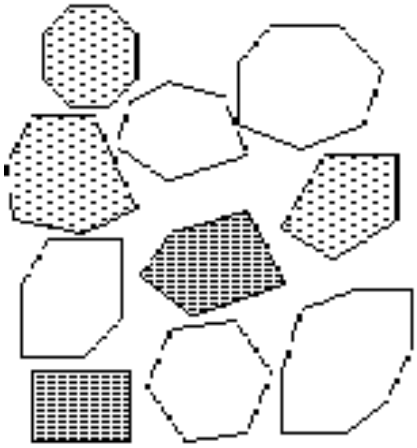
1. The participants examine the 50 polygons on the Collection of Polygons Activity Sheet and, working in pairs, choose an attribute that they can use to categorize the polygons.
2. Participants cut apart and place the polygons in sets based on the attribute that they chose. All polygons must belong to one category according to the attribute.
3. The instructor describes and illustrates the characteristics of a well-constructed bar graph.
    - A bar graph is a graph of data in which parallel bars are used for comparing information from several categories. Each bar represents one category and the length of the bar represents the number of times that category occurs.
    - The categories are represented on one axis of the graph by the bases of the bars touching the axis. Each of the bars, extending from the axis, represents a category and each bar is the same width. Equal empty spaces are left between bars. An equal empty space is left before the first bar and after the last bar.
    - The number count of how many are in a category is represented by a scale on the other axis. This axis is marked off beginning with a number lower than the lowest count of items in any category, not necessarily zero, and continuing to or above the highest count. Lines may extend from this axis behind the bars to make them easier to read.
    - The category axis should be labeled to describe the categories being counted and the numerical count axis should be labeled telling what the numbers represents.
    - The graph should have a descriptive title.
    - The instructor's example has all the attributes described and shown in the transparency.

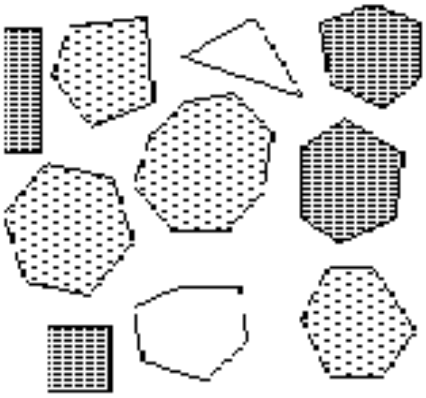
4. The participants work in pairs to construct a bar graph using the attribute that they chose.
5. When the pairs of participants have completed their graphs, they share them with other participants and assess whether or not they have included all the attributes of a well-constructed bar graph.
6. Each pair of participants chooses a different attribute of the set of polygons and constructs another bar graph based on the newly-chosen attribute.
7. When the pairs of participants have completed their second bar graph, they share them with another pair of participants and assess how well their graphs meet the attributes of a well-constructed bar graph.
8. Each participant chooses another attribute and constructs a bar graph independently. The bar graph should include all the attributes of a well-constructed bar graph. Below is one example of a graph participants may construct:

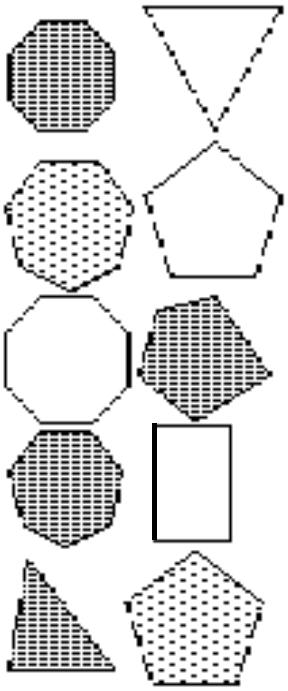
Shading of Polygons

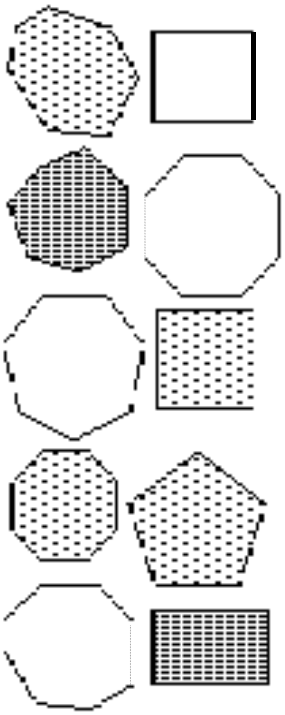


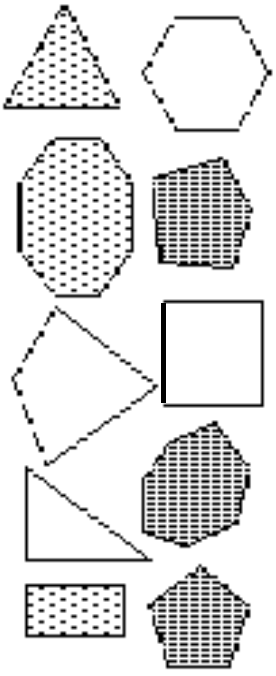
Collection of Polygons

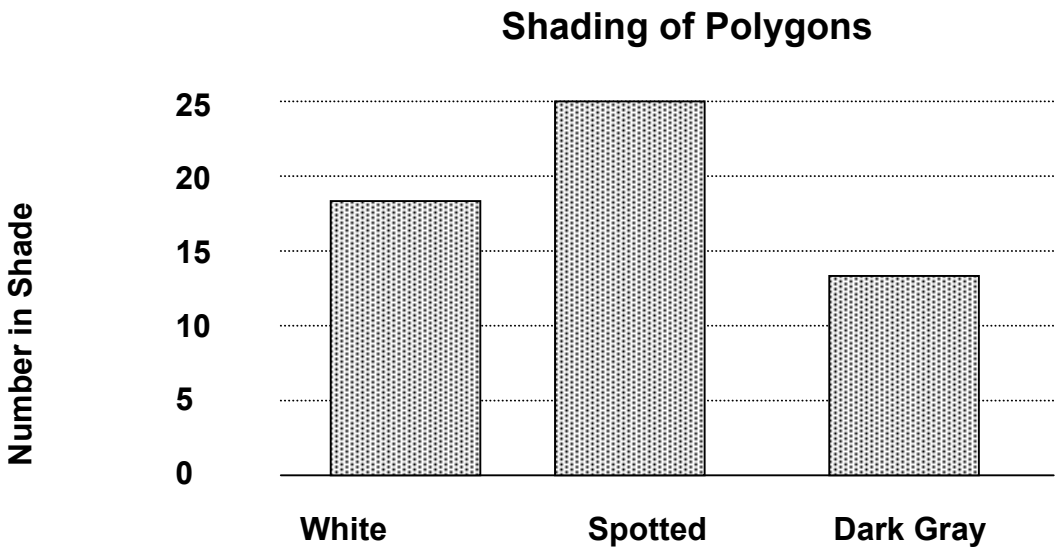











Shading of Polygons
Number in Shade
25
20
15
10
5
0
White
Spotted
Dark Gray